

RESEARCH ARTICLE

# Faecal Metaproteomic Analysis Reveals a Personalized and Stable Functional Microbiome and Limited Effects of a Probiotic Intervention in Adults

Carolyn A. Kolmeder<sup>1\*</sup>, Jarkko Salojärvi<sup>1</sup>, Jarmo Ritari<sup>1</sup>, Mark de Been<sup>1,2</sup>, Jeroen Raes<sup>3,4</sup>, Gwen Falony<sup>3,4</sup>, Sara Vieira-Silva<sup>3,4</sup>, Riina A. Kekkonen<sup>5</sup>, Garry L. Corthals<sup>6a</sup>, Airi Palva<sup>1</sup>, Anne Salonen<sup>1,7</sup>, Willem M. de Vos<sup>1,7,8</sup>

**1** Department of Veterinary Biosciences, University of Helsinki, Helsinki, Finland, **2** Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands, **3** KU Leuven, Department of Microbiology and Immunology, Rega Institute, Leuven, Belgium, **4** VIB, Center for the Biology of Disease, Leuven, Belgium, **5** Valio Oy, Helsinki, Finland, **6** Translational Proteomics, Turku Center for Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland, **7** Department of Bacteriology and Immunology, Immunobiology Research Program, University of Helsinki, Helsinki, Finland, **8** Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands

✉ Current address: Van't Hoff Institute for Molecular Sciences (HIMS), University of Amsterdam, Amsterdam, The Netherlands

\* [carolin.kolmeder@helsinki.fi](mailto:carolin.kolmeder@helsinki.fi)



## OPEN ACCESS

**Citation:** Kolmeder CA, Salojärvi J, Ritari J, de Been M, Raes J, Falony G, et al. (2016) Faecal Metaproteomic Analysis Reveals a Personalized and Stable Functional Microbiome and Limited Effects of a Probiotic Intervention in Adults. PLoS ONE 11(4): e0153294. doi:10.1371/journal.pone.0153294

**Editor:** Bryan A White, University of Illinois, UNITED STATES

**Received:** February 5, 2015

**Accepted:** March 28, 2016

**Published:** April 12, 2016

**Copyright:** © 2016 Kolmeder et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** Support was provided by: Doctoral Programme in Food Chain and Health (to CK); Academy of Finland [<http://www.aka.fi/en-GB/A/>] (141130 to WdV, 137389 to WdV, 141140 to WdV); European Research Council [<http://erc.europa.eu/>] (250172 to WdV); FP7 [[http://cordis.europa.eu/fp7/home\\_en.html](http://cordis.europa.eu/fp7/home_en.html)]; Spinoza award of the Netherlands Organization for Scientific Research [<http://www.nwo.nl/en/research-and-results/programmes/spinoza+prize>] (to WdV); METACARDIS HEALTH-F4-2012-

## Abstract

Recent metagenomic studies have demonstrated that the overall functional potential of the intestinal microbiome is rather conserved between healthy individuals. Here we assessed the biological processes undertaken in-vivo by microbes and the host in the intestinal tract by conducting a metaproteome analysis from a total of 48 faecal samples of 16 healthy adults participating in a placebo-controlled probiotic intervention trial. Half of the subjects received placebo and the other half consumed *Lactobacillus rhamnosus* GG for three weeks ( $10^{10}$  cfu per day). Faecal samples were collected just before and at the end of the consumption phase as well as after a three-week follow-up period, and were processed for microbial composition and metaproteome analysis. A common core of shared microbial protein functions could be identified in all subjects. Furthermore, we observed marked differences in expressed proteins between subjects that resulted in the definition of a stable and personalized microbiome both at the mass-spectrometry-based proteome level and the functional level based on the KEGG pathway analysis. No significant changes in the metaproteome were attributable to the probiotic intervention. A detailed taxonomic assignment of peptides and comparison to phylogenetic microarray data made it possible to evaluate the activity of the main phyla as well as key species, including *Faecalibacterium prausnitzii*. Several correlations were identified between human and bacterial proteins. Proteins of the human host accounted for approximately 14% of the identified metaproteome and displayed variations both between and within individuals. The individually different human intestinal proteomes point to personalized host-microbiota interactions. Our findings indicate that

305312 (to J. Raes); IWT-SBO 100016 (to J. Raes); Marie Curie Actions FP7 People COFUND [[http://cordis.europa.eu/project/rcn/100377\\_en.html](http://cordis.europa.eu/project/rcn/100377_en.html)] (Proposal 267139 (acronym OMICS@VIB) to SVS); Research Foundation Flanders (FWO) [<http://www.fwo.be/en/>] (to J. Raes); The Flemish agency for Innovation by Science and Technology (IWT) [<http://www.flagera.eu/?q=iwt>] (to J. Raes); KU Leuven [<http://www.kuleuven.be/english>] (to J. Raes); The Rega Institute [<http://rega.kuleuven.be/>] (to J. Raes); Biocenter Finland, Turku Proteomics Facility [<http://www.biocenter.fi>] (to GC). One of the authors (RAK) is employed by a commercial company (Valio Ltd) and contributed to the design of the clinical trial.

**Competing Interests:** The work was partially supported by Valio Ltd, a manufacturer of probiotics; one of the authors (RAK) is employed by Valio Ltd. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials, and the content of this paper was neither influenced nor constrained by that fact.

analysis of the intestinal metaproteome can complement gene-based analysis and contributes to a thorough understanding of the activities of the microbiome and the relevant pathways in health and disease.

## Introduction

During the last decade, many phylogenetic and metagenomic studies have revealed the vast genetic diversity and coding capacity of the intestinal microbiome in human adults [1,2]. We all harbour hundreds of microbial species, each carrying several thousands of genes and hence the intestinal microbiome enhances our own coding capacity by at least a hundred fold. Intestinal microbes are involved in many crucial functions, such as metabolizing food components that escape uptake by the host, contributing to the degradation of mucus, and producing vitamins and a variety of metabolites, such as short chain fatty acids and other important signalling molecules. Overall, the interactions taking place between the host and intestinal microbes, as well as between the different species of microbes, may exert a profound impact on the host's health [3,4]. The type and amount of microbes contributing to these actions differ substantially between individuals which means that the phylogenetic composition can be considered as kind of a subjective fingerprint [5,6]. However, metagenomic studies have shown only limited differences between subjects. Strikingly, while our microbiota varies phylogenetically, metagenomic analyses have revealed that at the highest functional level, such as in KEGG pathway maps, the functional potential of the microbiome of healthy individuals remains very similar [7,8]. However, this does not imply that there is a similarity in the actual activity of the microbiota in vivo, because this is affected by a variety of factors, including diet, host genetics, and health status.

Two well-established functional approaches have been exploited to extend the genome-scale analysis to activity based profiling: metatranscriptome and metaproteome analyses. Recent faecal and ileal metatranscriptomic investigations have indicated that there are pronounced differences at the gene expression level in subjects with comparable metagenomes [9–11]. However, bacterial and archaeal transcripts have short half-lives, in the order of seconds to minutes and hence faecal RNA may not accurately capture the biological processes taking place in the intestinal tract. In contrast to mRNA, proteins are more stable and may provide an interpretation of activities along the whole digestive tract. In the absence of global enzymatic or physiological approaches, metaproteomics, the study of all proteins in an ecosystem, has been applied to elucidate the actual function of the intestinal microbiota [12]. This technique makes it possible to monitor the activities of the host and microbes at the same time, and thus to study their mutualistic interactions [13].

So far, intestinal metaproteomic studies have been carried out with 20 healthy individuals at a maximum of two time points [13–16]. Functional differences based on proteins have been studied in a dozen patients with strong perturbing factors such as Crohn's disease or antibiotic treatment, but no functional baseline for healthy subjects has ever been established [16–19]. In general, functional differences between healthy individuals and their variation over time have only been studied to a limited extent.

Here, we expand these earlier studies with a cohort of 16 healthy subjects who took part in a placebo-controlled probiotic intervention trial and characterized the dominant functionalities by conducting a comprehensive faecal metaproteome analysis. In this longitudinal trial, these subjects were followed for six weeks and a total of 48 different faecal samples were taken;

before, during and after the placebo or probiotic intervention, resulting in three analysis time points. These were used to identify the temporally stable proteins in each subject, i.e. the individual core, as well as the shared functions, i.e. the common core, to define a metaproteomic baseline in healthy subjects. Moreover, we addressed the question of whether the microbial composition can explain the microbial activities i.e. converging vs. diverging in terms of composition and function.

## Materials and Methods

### Ethical statement

The Ethics Committee of the Hospital District of Helsinki and Uusimaa approved the trial and its protocol (Ethical protocol no HUS 3577E0/05; [Fig 1](#)). The trial is not registered in an international trial register. Written consent was given by the study participants prior to the start of the study. Details of this trial have been published previously [[20–22](#)].

### Study Design

The analysed samples were part of a randomized, double blinded intervention study testing the effect of various probiotic strains on immune and metabolic functions in healthy adults [[21](#)]. We focused here on the intervention with the well-described strain, *Lactobacillus rhamnosus* GG (ATCC 53103).

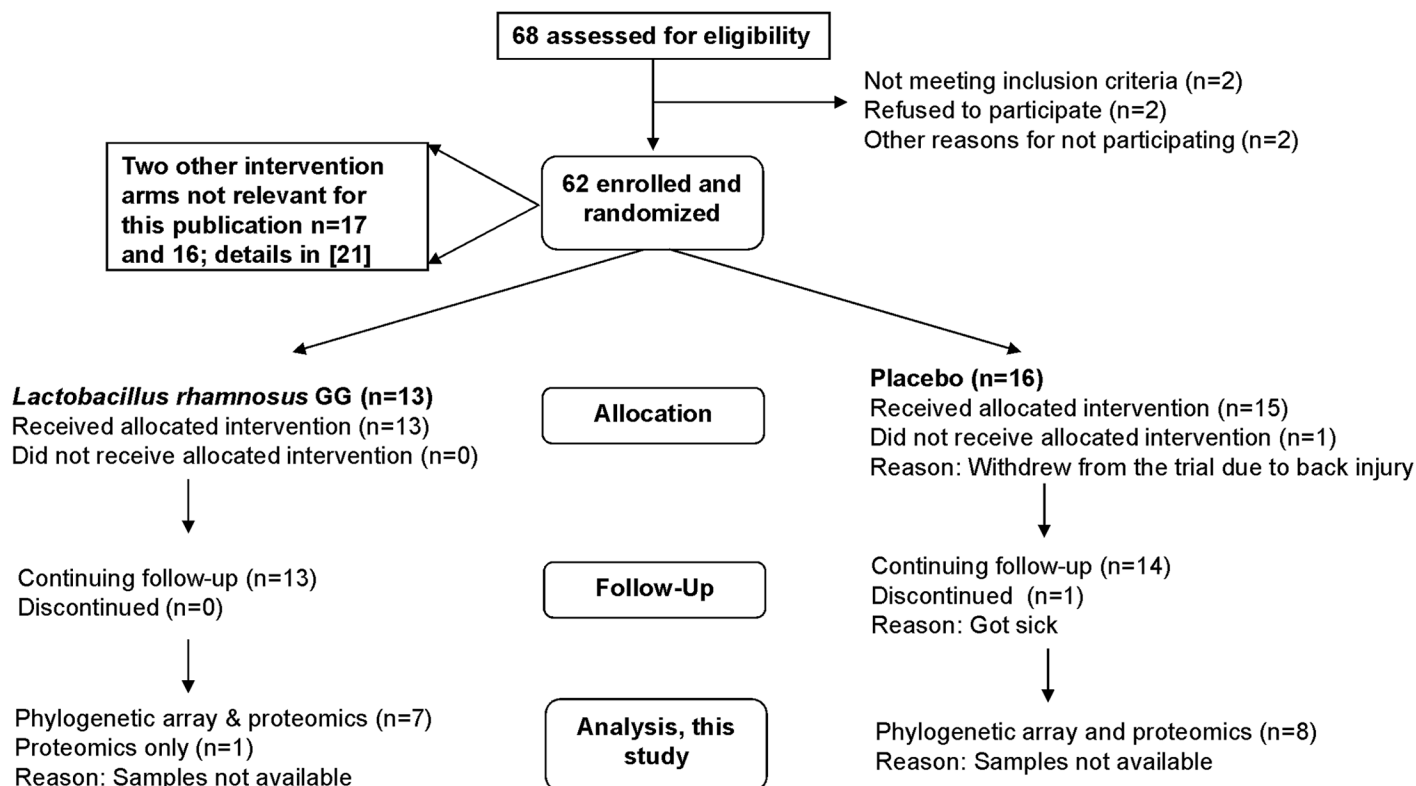
The trial consisted of a three-week run-in period, a three week intervention period and a three week wash-out period. Faecal samples at the end of each period were collected and stored at -20°C at home before transferring them within 24 hours to -80°C at the study centre. In this article, the sample collected at the end of the run-in period is called time point (TP) 1, the sample at the end of the intervention period is TP2, and the sample at the end of the wash-out period is designated TP3. During the intervention period study participants consumed each day 250 ml of a fruit based milk drink with or without *L. rhamnosus* GG ( $6.2 \times 10^7$  cfu/mL amounting to a daily dose of  $1.55 \times 10^{10}$  cfu). In the present analyses we selected samples for which faecal material was left. Eight subjects of the placebo group and eight from the *L. rhamnosus* GG intervention group were included ([S1 Table](#)). The reduction of the study subjects due to sample availability resulted in a skewed age distribution between the placebo and LGG group (medians of 30 and 49 years, respectively). The median body mass index (BMI) was 23.9 in the placebo group and 25.9 in the LGG group. The female:male ratio was 6:2 in the placebo group and 5:3 in the *L. rhamnosus* GG group. As detailed in [[21](#)], trial participants kept a study diary. These included reporting any intestinal symptoms, the use of antibiotics and the regular use of probiotics ([S2 Table](#)). All individuals rated their health status as good, except for subjects 116 and 163 who described their health status as average.

### Protein extraction

Proteins were mechanically extracted from faecal samples (125 mg) by bead beating essentially as previously described [[13](#)]. Six cycles of bead beating (6.5 ms<sup>-1</sup> for 45 seconds) were run on a FastPrep 24 (MP Biomedicals). The samples were kept on ice during the required 5 min resting period in-between the cycles.

### 1D gel electrophoresis and in-gel protein digestion

To reduce the complexity of the protein extract, the same 1D gel fractionation approach was carried out as described in [[13](#)]. The 37 kDa and the 75 kDa band of a prestained marker (Precision Plus™ Dual Color, Biorad) was used to define the height for cutting the lane into three



**Fig 1. Consort flowchart of trial with ethical protocol no HUS 3577E0/05.**

doi:10.1371/journal.pone.0153294.g001

fractions. Only the middle part (~37–75 kDa), containing the abundant metaproteome (AMP) was subsequently analysed. In order to control for possible run differences impairing reproducible cutting, one of the samples was run on all of the gels. Gel pieces were washed and proteins reduced, alkylated, and tryptically digested (per sample 200 ng of trypsin sequencing grade, Promega) overnight similarly as described in [23]. Following extraction from the gel-pieces, peptides were vacuum dried and stored at -20°C until used. Dried peptides were resolved in 2% formic acid and filtered through glass microfiber filters (Whatman; grade GF/C) for removal of gel particles. The peptide elutes were desalted by using OMIX tips (Agilent Technologies). After aspirating the samples 15 times, the peptides were eluted with 60% acetonitrile/2% formic acid, dried and reconstituted with 1% formic acid prior to LC-MS/MS.

### Liquid chromatography tandem mass spectrometry

Protein digests were analysed by LC-MS/MS on a nanoflow HPLC system (Easy-nLCII, Thermo Fisher Scientific) coupled to a LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nano-electrospray ionization source. The mobile phase consisted of water/acetonitrile (98:2 (v/v)) and 0.2% formic acid (solvent A) and acetonitrile/water (95:5 (v/v)) and 0.2% formic acid (solvent B). Ten µl of sample were injected into LC-MS/MS equipment. Peptides were loaded onto a reversed phase precolumn (in-house made, 2 cm long, 100 µm inner diameter) with solvent A and separated in a reversed phase analytical column (in-house made, 15 cm long, 75 µm inner diameter). Both columns were packed with Magic C18-AQ 5 µm 200 Å resin (Michrom BioResources Inc., Sacramento, CA,

USA). Peptides were eluted by using a step-wise gradient (2–20% solvent B for 70 minutes and 20–40% solvent B for 30 minutes at a flow rate of 300 nl/min).

Full scan for eluting peptides was acquired in the mass range of 300–2,000 m/z of the Orbitrap-detector with 30,000 resolution at 400 m/z, AGC target set to 1,000,000 and maximum inject time set to 100 ms. Based on the results of the full scan, the 20 most intense ions were fragmented by CID fragmentation (normalized collision energy of 35%, activation time of 10 ms and activation Q set to 0.25) and scanned on the LTQ linear ion trap (AGC target set to 5000 and maximum inject time set to 50 ms). An isolation width of 2 m/z was used for precursor selection. Precursors whose charge state could not be determined or for which the charge state was +1 were discarded from the MS/MS analysis. Precursors were dynamically excluded for 30 s with a repeat count of one. Both MS and MS/MS scans consisted of one microscan. MS scans were acquired as profile data and MS/MS scans as centroid data. The peaks were extracted with the software proteome discoverer (Thermo Scientific) with default settings and the resulting Mascot generic format (mgf) files are available at <ftp://MSV000079596@massive.ucsd.edu> (generic password “a”).

## Peptide identification

Peptide identification was performed as described previously [13] with a database extended for the number of bacterial genomes and alternative human proteins (S3 Table). Peptide identifications were filtered with a false discovery rate (FDR) of 5%.

## Functional and phylogenetic annotation of peptides

Functional annotation of peptides was performed by a BLAST search of the original protein sequences against the Cluster of Orthologous Groups database (COGdb) as in [13]. In addition, a BLASTP 2.2.27 search of the original protein sequences against NCBI nr was performed. In the KEGG Orthology (KO) retrieval, BLAST results were imported into the metagenomic data analysis software MEGAN [24] (version 5.1.5; default parameters: maxMatches = 100, minScore = 50.0, maxExpected = 0.01, topPercent = 10.0, minSupport = 50).

A possible phylogenetic origin was assigned to a peptide by the same principle, as described previously [13]. In brief, in order to achieve a successful taxonomic annotation a peptide was required to have a complete peptide match in the universal protein resource knowledgebase (UniProtKB) and in the case of multiple taxonomic hits per peptide the lowest common ancestor was reported. Instead of utilizing in-house Python scripts [13] the online tool unipept was used [25] (version 1.4; settings: I and L equated, advanced missed cleavage option (trypsinized peptides searched but results per miss-cleaved peptide combined), duplicates filtered).

In the quantitative analysis, spectra per peptide per class (COG, KEGG, phylogenetic level etc.) were summed and relative percentages per sample calculated for heatmap visualization.

To assess the contribution of distinct functionalities in a certain domain, genus, or species, the functionally and phylogenetically annotated spectra were queried for the taxonomic unit of interest. To create a subset of the data containing peptides with a human origin and to investigate the human functionalities in more detail, peptides with at least “Chordata” assignments were picked. Selecting only *Homo sapiens* specific peptides would have resulted in very few hits, due to the wide range of Chordata species with highly similar homologous sequences. Bacterial KOs were mapped onto microbial gut-specific metabolic modules (GMM), its syntax is provided in the supplement in [26]. The GMM presence is defined with a threshold of detection of more than two thirds of the KOs involved in one of the alternative metabolic pathways covered by a module (i.e. coverage > 2/3). The GMM abundance is defined as the median KO abundance in the maximum coverage alternative pathway, if coverage > 2/3.



## Rarefaction analysis

KO counts for the lowest functional level (i.e. fourth) were used for analysing how the numbers of found KOs change as a function of number of samples. All samples per individual were combined and the results of a rarefaction analysis including all possible combinations both for accumulated (KO rarefaction) and shared KOs (KO core) were plotted. The analysis was performed at intervals ranging from two to 16 samples. At each interval, 1000 bootstraps were taken from the samples, and the numbers of KOs present in one or more of the samples included in a bootstrap were counted. A similar analysis was performed by counting the number of KOs present in all of the samples included in a bootstrap. This resulted in an estimation of sampling distribution of KO counts at varying levels of sample numbers. The analysis was performed with R v.3.1.1.

## Phylogenetic analysis: microarray and quantitative polymerase chain reaction (qPCR)

Extraction of DNA was performed with the Promega kit; this involved homogenization of the faecal material with ample amount of 1x PBS and continuing DNA extraction with the pellet as described in [27]. A phylogenetic microarray (Agilent Technologies), the Human Intestinal Tract Chip (HITChip), was applied to analyse the phylogenetic composition based on 16S rRNA gene amplicons [28]. Hybridisation of the microarray and min-max signal normalization were performed as described [29]. Per target sequence the oligonucleotide probe signals were summed [30]. Throughout this study, relative abundances are provided at either the phylum- or the genus-like level. The genus-like level microarray data are provided in S4 Table. In addition to microarray analysis, qPCR of 16S rRNA regions was carried out to determine the total amount of bacteria, Bacteroidetes, *Bifidobacterium* spp. and *Lactobacillus* spp. present in 1g of faeces using taxa-specific primers as described previously [31]. The presence of archaeal DNA was analysed with qPCR, using primers specific for *Methanobrevibacter* [32].

## Statistical data analysis

The statistical programming language R was used for statistical data analysis [33]. Hierarchical clustering using complete linkage algorithm and the Jaccard distance between identified peptides was used to determine the dissimilarity of the metaproteome [34]. Hierarchical clustering (complete linkage) was also applied to the oligonucleotide profiles using 1-Pearson's correlation coefficient ( $r$ ) as the similarity measure.

Split violin plots (R package vioplot and <http://mbjoseph.github.io/blog/2013/06/24/violin/>) were used to visualize the distribution of relative abundances of bacterial phyla both for phylogenetic microarray and proteomic data.

To identify correlations between human and bacterial functional classes, Spearman correlation was applied, and the R package qvalue was used for calculating false discovery rate adjusted p-values (q-values). Results were filtered for a q-value < 0.05.

## Results

### Microbial and human proteins in the cohort reveal specific core functions

The analysis of the metaproteome of 48 faecal samples obtained at the three time points, from each of the 16 subjects, which had been selected from the probiotic and placebo groups (Fig 1), resulted in approximately 1 million mass spectra of which 213,855 (23.7%) could be identified (FDR < 5%) (S5 Table). Because of the ambiguity in mapping the peptides to distinct proteins—the so-called protein inference—that is especially noticeable in metaproteomics, we have

compared samples based on their spectral counts per functional group and not per protein [12,35].

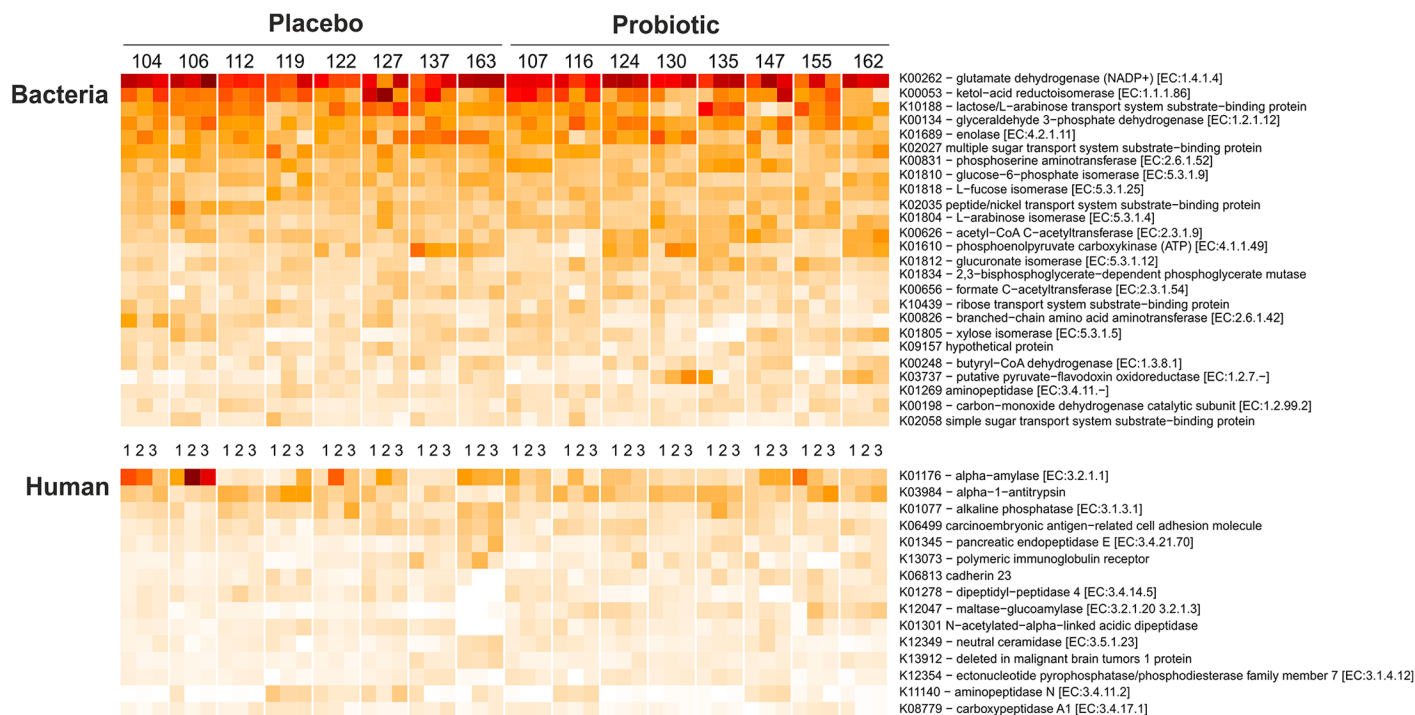
Functional annotations provided a COG or KO assignment for 78% of the identified spectra. In addition, we could assign a taxonomic origin according to the lowest common ancestor principle to 66% of the identified peptides, of which 80.9% were classified as bacterial, less than 1% as archaeal, and 13.8% as eukaryotic. The remaining 5.3% could not be assigned to a single domain as the peptides could be found in at least two domains of life. The differentiation into bacterial and eukaryotic/human peptides was utilized for determining domain-specific functionalities.

Overall, 967 different bacterial KOs were identified. Rarefaction analysis of the functional diversity across the 16 individuals (S1 Fig) showed that more than half of these bacterial KOs were already covered when combining the analysis of three individuals. A core of around 200 KOs was identified in any pair of two subjects and around 100 KOs were found in each of the individuals.

It was decided to concentrate on relevant intestinal functions and focus on those proteins present in the highest amounts. Therefore, we chose those KOs for which at least five spectra in one of the samples across the 48 samples had been identified (S6 Table). From these 238 remaining KOs, the 25 most abundant bacterial ones were identified in all of the samples, representing the most stable and abundant core functions (Fig 2). Glutamate dehydrogenase was found as the most abundant intestinal functionality (on average 7% of the spectra) in healthy individuals, as reported previously [13]. Other abundant proteins included major glycolytic enzymes as well as specific sugar-converting enzymes such as L-fucose isomerase, L-arabinose isomerase and glucuronate isomerase. The 238 bacterial KOs were mapped onto microbial gut-specific metabolic modules (GMM). A total of 21 out of 102 modules was detected as being enriched (S7 Table). With an average enrichment of 38 spectra the arabinose module was found to be extensively represented in all subjects at all time points. With respect to the COG family level, the metaproteomes appeared to be very similar (Fig 3). The COG families for carbohydrate transport and metabolism, energy production and conversion and amino acid transport and metabolism were the most representative families in all of the samples. Despite the rather even distribution of COG families across the samples, individual patterns were clearly discernible. These included a relatively higher level of COG family L (Replication, recombination and repair) in the samples of subjects 106, 137 and 124. Moreover, a clear difference could be observed in those peptides to which no COG family could be assigned, highlighting the importance of further studying these unassigned hits.

A total of 134 human KOs was identified. We selected those KOs with at least two spectra in one of the samples across the 48 samples, which resulted into 73 KOs (S8 Table). From these, the most abundant 15 KOs represented functions involved in digestion and host protection, and these showed specific patterns among all subjects and time points (Fig 2). Proteases and peptidases accounted for an average of 23% of the 73 most abundant human KOs. Carbohydrate and lipid-degrading enzymes were less abundant.

Correlation analysis of the human and bacterial KOs revealed 138 potentially relevant correlations (p-value and q-value < 0.05, S2 Fig). These included significant positive correlations between the human interacting serine/threonine-protein kinase 4 (K08848) and the bacterial flagellar hook protein FlgE (K02390), i.e. evidence for the presence of flagellins; between human alpha-amylase (K01176) and bacterial putative ABC transport system substrate-binding protein (K01989), propionate CoA-transferase (K01026), ATP-dependent Clp protease ATP-binding subunit ClpB (K03695), and cytidylate kinase (K00945), as well as a negative correlation between the antimicrobial protein DMBT1 (K13912) and the bacterial bleomycin hydrolase (K01372). Another indication for specific host-microbiota interactions may be



**Fig 2. Top 25 bacterial (upper panel) and top 15 human (lower panel) functional categories (KEGG).** (white to dark red range for bacteria KOs are from 0.1 to 10.2% out of 238 bacterial KOs and for human KOs from 0.0 to 49.3% out of 73 human KOs). The samples are organized according to treatment group (left: placebo group, right: probiotic group); sample time points are abbreviated as 1 (sample collected at the end of the run-in period), 2 (sample collected at the end of the intervention period) and 3 (sample collected at the end of the wash-out period).

doi:10.1371/journal.pone.0153294.g002

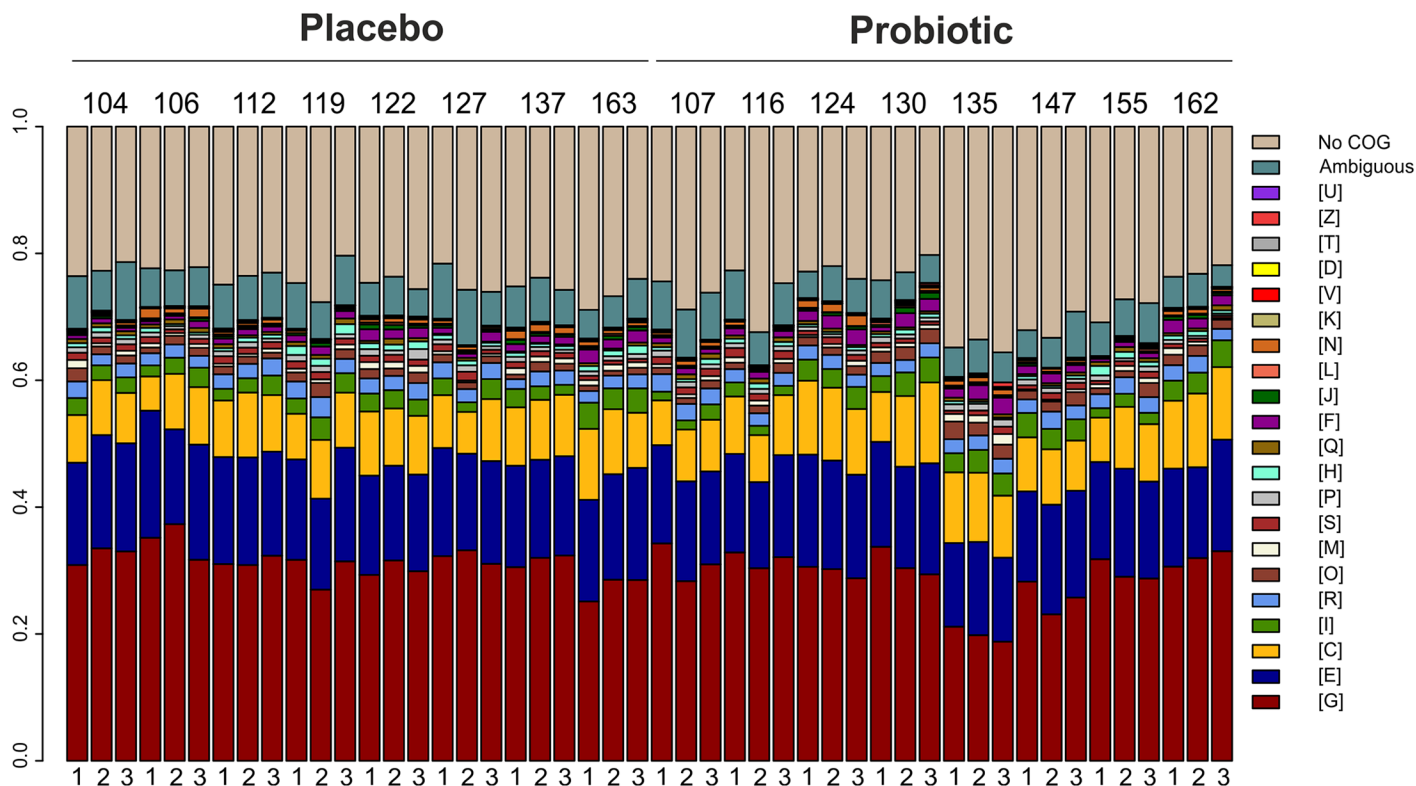
deduced from the identification of an enrichment of degradation of the host-derived glycan mucin in at least two samples of four subjects (107, 119, 147, and 155) when mapping the bacterial KOs onto GMM.

## Evidence for a personalised and stable functional microbiome

A subject-specific clustering of all 16 subjects was observed when the identified peptides were used for hierarchical clustering based on Jaccard similarity (Fig 4). Moreover, also based on KO abundances (at least 5 spectra in one sample), virtually all of the samples clustered according to the individual subject (S3 Fig). This indicates that the functional microbiome is personalised and represents an extension of the compositional individuality that has been described previously and is now confirmed here (S4 Fig). Similarly, when comparing the peptide-based and the phylogenetic oligonucleotide-based hierarchical clustering in more detail, we observed that the samples obtained from subject 135 were quite distant from the other samples. However, in the peptide hierarchical clustering the samples of subject 119 were even more different than those of subject 135. This indicates that the proteins of subject 119 are rather unique and may reflect an aberrant health status and/or antibiotic use, as reported by this subject (S2 Table).

To compare the individual metaproteomes on a functional level, we studied the individual differences of certain gut-specific KOs (Fig 5). The proteins involved in the production of short chain fatty acids (SCFA) are of interest due to the multitude of beneficial roles that SCFAs play in the intestine. The average levels for the KO for butyryl-CoA dehydrogenase, an enzyme involved in the production of butyrate, varied from 0.2% to 1.4% in the subjects. For





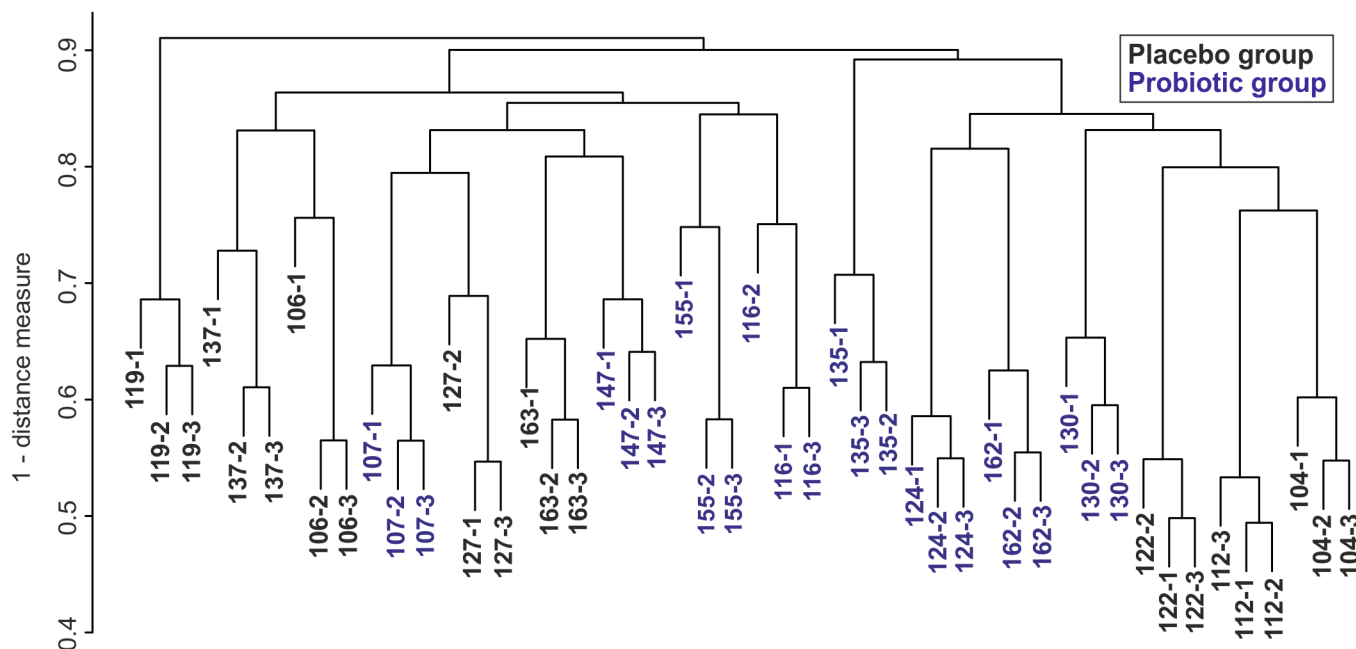
**Fig 3. Relative distribution of COG families across samples.** The samples are organized according to treatment group (left: placebo group, right: probiotic group); sample time points are abbreviated as 1 (sample collected at the end of the run-in period), 2 (sample collected at the end of the intervention period) and 3 (sample collected at the end of the wash-out period); ambiguous = functions being part of more than one COG family; [G] Carbohydrate transport and metabolism; [E] Amino acid transport and metabolism; [C] Energy production and conversion; [I] Lipid transport and metabolism; [R] General function prediction only; [O] Post-translational modification, protein turnover, and chaperones; [M] Cell wall/membrane/envelope biogenesis; [S] Function unknown; [P] Inorganic ion transport and metabolism; [H] Coenzyme transport and metabolism; [Q] Secondary metabolites biosynthesis, transport, and catabolism; [F] Nucleotide transport and metabolism; [J] Translation, ribosomal structure and biogenesis; [L] Replication, recombination and repair; [N] Nucleotide transport and metabolism; [K] Transcription; [V] Defense mechanisms; [D] Cell cycle control, cell division, chromosome partitioning; [T] Signal transduction mechanisms; [Z] Cytoskeleton; [U] Intracellular trafficking, secretion, and vesicular transport.

doi:10.1371/journal.pone.0153294.g003

methylmalonyl CoA mutase, a key enzyme in propionate synthesis, levels varied from less than 0.1% to 1.2%. In general, the levels of KOs involved in butyrate synthesis were higher than those for propionate (with butyryl-CoA dehydrogenase plus 3-hydroxybutyryl-CoA dehydrogenase and methylmalonyl-CoA mutase as proxies). For two subjects (119 and 155), similar amounts of both pathways were found (S5 Fig). Flagellins are signalling ligands of the toll like receptor (TLR) 5 and thus they represent microbial proteins that communicate with the host. This class of protein varied in the subjects from 0.2% to 1.6% (Fig 5).

The predominantly anoxic environment in the colon and other parts of the intestinal tract requires that the intestinal microbiota have to exploit anaerobic strategies for dealing with electrons. It was predicted that pyruvate-flavodoxin oxidoreductase might contribute to an anaerobic electron sink and this protein class showed a distribution varying from 0.4% to 2.9%. Sirohydrochlorin cobaltochelate, an enzyme involved in the anaerobic biosynthesis of cobalamin (vitamin B12), contributed between 0.1% and 1.0% to the 238 most abundant KOs.

The consumption of LGG did not exert any effect on the identified peptides nor on the KEGG profiles (Fig 2; S3 Fig) i.e. the probiotic intervention did not lead to a systematic change of the identified peptides and their associated function. Since we tested this hypothesis with different statistical methods, we are confident that we did not detect functional differences. When



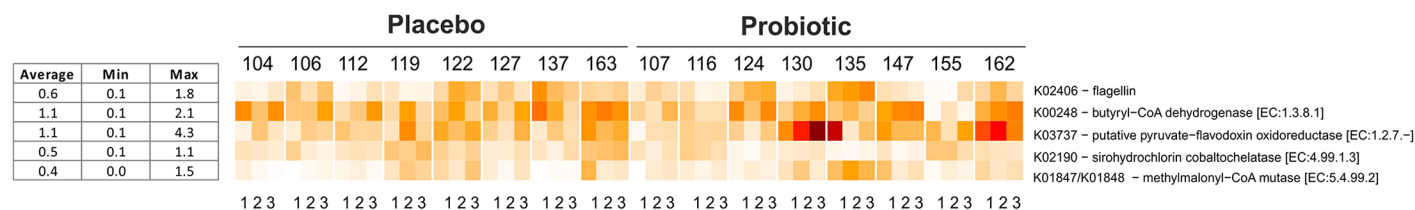
**Fig 4. Hierarchical clustering of Jaccard similarity indexes of identified peptides.** The peptide profiles clustered according to each subject however not according to the treatment group. The samples from subject 119 differed the most and 135 the next most from the rest of the samples (at a distance measure of 0.9 and just below, respectively); labels: Subject—time point; 1: sample collected at the end of the run-in period, 2: sample collected at the end of the intervention period, and 3: sample collected at the end of the wash-out period; black font: placebo group, blue font: probiotic group.

doi:10.1371/journal.pone.0153294.g004

calculating the ratio of bacterial to eukaryotic assignments, we did notice an increase during and a decrease after the intervention in the LGG group such that the samples from the LGG group clustered significantly together (S6 Fig; Pearson correlation; Fisher exact test p-value: 0.04). While this may point towards a decreased level of faecal host proteins in the LGG intervention group, stringent statistical analysis between the ratio of bacterial/eukaryotic peptides at time point one versus time point two in the LGG group did not reach statistical significance ( $p = 0.08$ ).

### Variation in human proteins

The three most abundant human KOs were the starch-degrading enzyme alpha-amylase (average 11.7%), the serpin alpha 1 antitrypsin (10%) and the hydrolase alkaline phosphatase (6.1%); these revealed differences across and within individuals (Fig 2). The levels of alpha-amylase differed substantially between individuals. Relative abundances reached almost 50% at



**Fig 5. Heatmap of KEGG functions with marked differences between subjects.** (white to dark red range for bacteria KOs are from 0.0–4.3% out of 238 bacterial KOs). The samples are organized according to treatment group (left: placebo group, right: probiotic group); sample time points are abbreviated as 1 (sample collected at the end of the run-in period), 2 (sample collected at the end of the intervention period) and 3 (sample collected at the end of the wash-out period).

doi:10.1371/journal.pone.0153294.g005

TP2 and TP3 in subject 106, but were consistently low (4–6%) in subjects 135 and 137. Notably, low alpha-amylase levels co-occurred with high levels of *Prevotella* spp. (see below).

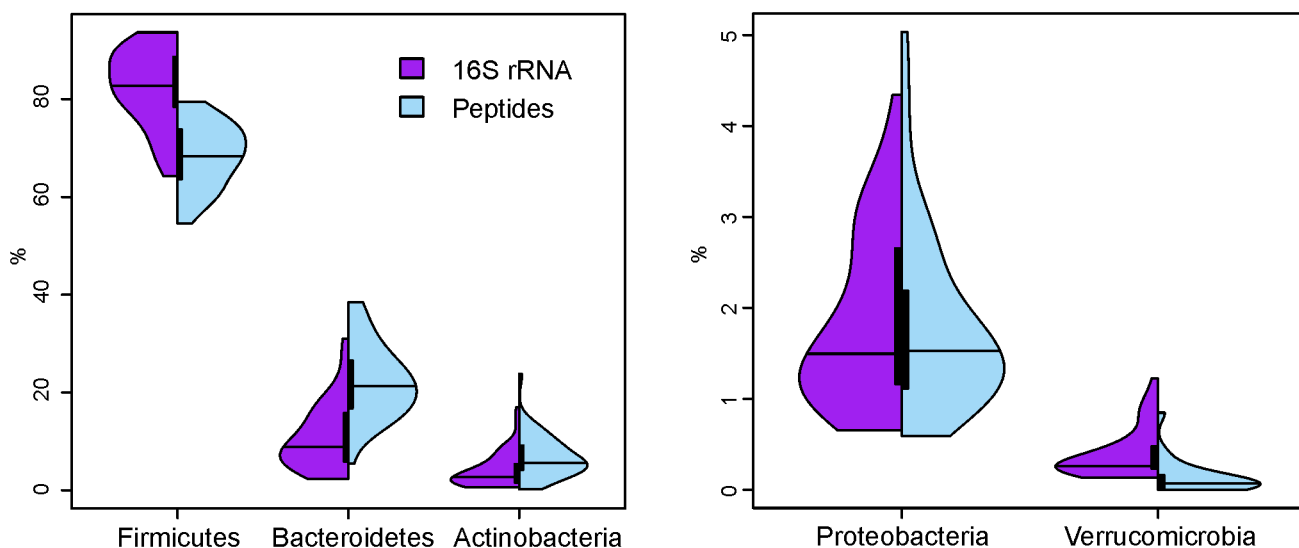
High levels of alpha 1 antitrypsin of 10% were found in the majority of the samples with a minimum of 4.5% and a maximum of 18.4% of the human spectra.

## Phylum origin of 16S rRNA genes and peptides

A phylogenetic 16S rRNA gene based microarray (HITChip) was used to assess the microbial composition of the studied faecal samples. The average diversity (Shannon index) was 5.89 (min: 5.54, max: 6.3) in line with normal diversity observed in adults as observed in the in-house HITChip database with more than 1000 adult samples, which has been described previously [36]. The oligonucleotide profiles clustered according to subjects (S2 Fig). However, temporal variation was observed both at the phylum- and genus-like phylogenetic levels (S9 Table).

On average, more than 99% of 16S rRNA gene signals and MS/MS spectra with microbial assignment could be assigned to the four phyla Firmicutes, Bacteroidetes, Actinobacteria, and Proteobacteria in descending abundance (Fig 6; data per sample in S7 Fig). The three main phyla varied extensively throughout all samples for both measurements. For example, Actinobacteria varied from below 1% to over 15% both for 16S rRNA gene and proteomic measurements. When compared to the 16S rRNA gene data, there were less Firmicutes MS/MS assignments but more Actinobacteria and Bacteroidetes MS/MS assigned. The distribution of Proteobacteria was rather similar with the HITChip and proteome data. Verrucomicrobia, the fifth most abundant phylum, had a slightly higher representation on 16S rRNA gene than at the peptide level. Looking at the ranking of the phylum the microarray and metaproteomic analyses agreed with the order of Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, and Verrucomicrobia.

A high correlation (Pearson correlation coefficient of 0.67 at a q-value below 0.05) was found between phylogenetic microarray signals and spectral counts for Actinobacteria.



**Fig 6. Distribution of the five main phyla in the studied cohort on 16S rRNA and peptide level.** The split violin plots present side by side the relative amount (in percentage) of Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria and Verrucomicrobia obtained by 16S rRNA gene analysis (violet) and the metaproteomic measurements (blue). Both distributions are based on all 44 samples for which both 16S rRNA and metaproteomics data were available. The lines denote the median of the distributions (based on all the measured samples with the two different techniques) and the bar represents the 0.25 to 0.75 quantile.

doi:10.1371/journal.pone.0153294.g006

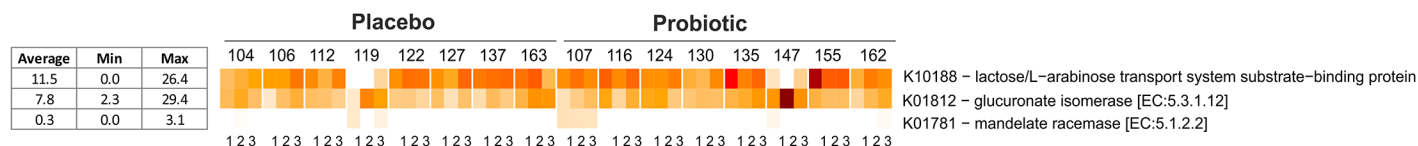
Bacteroidetes and Firmicutes correlated negatively, at the phylogenetic microarray level as well as at the peptide level. However, in general, the Bacteroidetes peptide assignments were higher than the 16S rRNA gene measurements and only in a few cases was the phylogenetic array signal higher than the ratio of assigned peptides (S7 Fig). In addition, the Bacteroidetes microarray signals displayed more differences between the three time points, whereas the Bacteroidetes levels determined by taxonomic assignment of the identified peptides were rather stable. To control for possible bias originating from the microarray measurements, Bacteroidetes qPCR was carried out and these results confirmed the phylogenetic microarray data. The abundance of the most common intestinal methanogen i.e. the Euryarchaeon *Methanobrevibacter smithii*, was estimated by performing qPCRs; in 12 of the 44 samples more than  $10^8$  16S rRNA gene copies/g of faeces were detected. For five of these 12 samples, the Euryarchaea-specific peptide assignments were higher than 0.1% of the phyla assignments. In general, within 10,000 peptides with an assigned taxonomy, only about four Euryarchaea peptides were identified, with the exception of sample 135–3 in which almost 1% of peptides with a phylum assignment were of Euryarchaeal origin. Fungal peptides were identified at a similarly low level and around 4 out of 1000 spectra with taxonomic information were found to derive from plants (data not shown).

Plotting the COG assignments per phylum revealed a similar pattern as that previously observed with a smaller set of samples (S8 Fig) [13]. More than 59% of the Actinobacteria spectra were grouped to the category of Carbohydrate Metabolism, highlighting the impact of that phylum on sugar utilization and transport. Almost 50% of the Bacteroidetes spectra had no assigned function, emphasizing the further need for bacterial physiological studies to explore these unknown encoded functions.

## Pathways of prominent and distinct gut microbes

The most prominent microbial species in general (on average constituting 15% of the microbiota measured by 16S rRNA gene analysis) was *Ruminococcus obeum*, a well-known representative of the largest phylum Firmicutes. However, on average, only 1.6% of the bacterial peptide hits classified at least to a phylum assignment were specifically assigned to *R. obeum*, hampering comparisons between samples. In contrast, the second most abundant species based on phylogenetic data, *Faecalibacterium prausnitzii* (constituting around 10% of the microbiota assessed by microarray measurements), represented on average 9.8% of the spectra. Therefore, we further analysed the distribution of *F. prausnitzii* and its activities. Bacteria related to *F. prausnitzii* showed substantial variation across subjects and time points at the 16S rRNA gene level (S9 Fig). For some subjects, 16S rRNA gene and proteomics measurement displayed essentially the same trend (e.g. samples of subjects 104, 106, 112) but for some samples, the 16S rRNA gene and peptide measurements developed in the opposite directions. In particular, subject 137 was a standout example where *F. prausnitzii* was found to contribute less than 5% to the overall community, whereas its metaproteome exceeded more than 10% of all proteins, indicative of the very high gene expression activity of this bacterial group. The majority of the *F. prausnitzii* KOs were involved in amino acid metabolism, but almost half of the *F. prausnitzii* peptides could not be designated to any specific function. The analysis of *F. prausnitzii*'s functional variation showed a certain trend within individuals (Fig 7). For example, glucuronate-isomerase was detected in relatively high amounts across all samples with few exceptions. In subjects 119 and 147, very low levels of lactose/L-arabinose transport system substrate-binding protein were observed.

For the second largest phylum Bacteroidetes species-specific spectra were much less abundant than for the above-mentioned dominant Firmicutes. This was also found for the spectra



**Fig 7. Heatmap of selected KOs assigned to *F. prausnitzii*.** (white to dark red range for 0.0–29.4% out of 123 *F. prausnitzii* specific KOs). The samples are organized according to treatment group (left: placebo group, right: probiotic group); sample time points are abbreviated as 1 (sample collected at the end of the run-in period), 2 (sample collected at the end of the intervention period) and 3 (sample collected at the end of the wash-out period).

doi:10.1371/journal.pone.0153294.g007

of *Prevotella* spp., which in contrast to *Bacteroides*, are believed to be more prevalent in individuals consuming fiber-rich diets. For example, *Prevotella* methylmalonyl-CoA mutase was found in all of the samples of subject 162, indicating stable *Prevotella*-specific propionate synthesis in that individual (S10 Table). *Prevotella* microarray signals were around the mean of 1% but for subjects 130 and 135 they were up to 20%. For all the samples of these two individuals, several hundreds of *Prevotella*-specific peptides were identified especially for carbohydrate degrading enzymes (S10 Table). In addition, low levels of human alpha-amylase were found in the samples of these individuals. This may hint that high amounts of complex carbohydrates were part of these individuals' diets and left for *Prevotella* to degrade.

Since Actinobacteria, the third largest phylum, have a distinctive lifestyle in the intestine, i.e. specialization in the utilization of complex carbohydrates, we attempted to clarify whether there would be any detectable intra-individual differences within this phylum. We found that Actinobacterial xylose isomerase was missing from the samples from individuals 119 and 135. There was variation in the amounts of Actinobacterial beta-fructofuranosidase between individuals, for example, low levels were detected in subjects 107 and 135 (S10 Table).

Some differences were found in the abundances of Proteobacteria KO, the fourth largest phylum in this study. Proteobacterial thioredoxin 1 was found in all of the samples but at a very low level. A subject-dependent pattern was found for Proteobacterial taurine-pyruvate aminotransferase (S10 Table). This enzyme mostly originated from *Bilophila*, a genus found to be present in a large proportion of the population and which produces sulphide from taurine [37]. Taurine has multiple roles in humans, such as the formation of bile salts, and *Bilophila* spp. may be involved in a balanced turn-over [38].

In the samples with relatively high levels of archaeal proteins (from subjects 135-2/3, 162-3) KOs involved in methane metabolism were detected, most prominently the methyl-coenzyme M reductase beta subunit (S11 Table).

## Discussion

Metaproteome analysis is a recently developed approach to identify at the protein level the highly expressed genes in an ecosystem [12,39].

Here, we studied the faecal metaproteome of 16 healthy Finnish adults who took part in a probiotic intervention and consumed a fruit-based drink with or without *L. rhamnosus* GG by sampling at three time points over a period of six weeks and found that the functional microbiome is both personalized and stable over time. Addressing bacterial, archaeal and human proteins, we were able to provide a state of the art overview of the activities of both the intestinal microbiota and the host.

## Functionality of the gut ecosystem

Around 10% of the identified microbial KO functions were found to be present in all the individuals; these represent the core of the intestinal ecosystem with carbohydrate transport and



metabolism, energy production and conversion and amino acid transport and metabolism being the most prominent metabolic routes. Obviously, by analysing the metaproteome of 16 subjects there is not yet a saturation of identified functions but by adding more individuals more functions will still be identified. The identification of several sugar-converting enzymes in the core metaproteome is indicative of the microbiome's ability to exploit glycans as a rapidly available energy supply. The observation that the arabinose module was found to be enriched in all individuals might be attributable to the fact that Finnish subjects are known to consume large amounts of fibres, specifically those derived from rye, which also contains arabinose [40–42].

Similar to the microbial core functions, also a core of the host functions could be detected. The most prominent biological activities found in the host core were proteinases and peptidases, indicative of high intestinal proteolytic activity or the high stability of these proteins, since only free amino acids and di- and tri-peptides are transmitted to the blood stream in the intestine. However, these proteinases and peptidases may also be acting on different routes and may possess signalling activities.

Other host core proteins included important functions such as the starch-degrading enzyme alpha-amylase (average 11.7%), the serpin alpha 1 antitrypsin (10%) and the hydrolase alkaline phosphatase (6.1%); these enzymes displayed differences not only between but also within the same individuals (Fig 2).

## Personalized and stable functional intestinal microbiome

We found that each adult contained a unique faecal metaproteome that was stable over the analysed period of six weeks. Therefore, compositional individuality was conserved at the functional level, suggesting that a distinctive set of microbes may be responsible for a specific distribution of functions. Further studies will be required to identify the factors involved. Most likely, host genes together with environmental factors (diet, drugs, hygiene setting) may prime the ecosystem to develop its individuality. The presence of a personalized functional microbiome was suspected based on peptide assignments and KO abundances. Thus, there seems to be a functional stable metaproteome and sequence differences (as detected at the peptide level) due to different microbial communities may not solely explain the clustering of the subjects. A comparison of the clustering at the oligonucleotide and peptide level could suggest that there is stronger individuality at the phylogenetic level in comparison to the metaproteomic data. However, this observation can be partly explained by the difference in the type of measurement and the subsequent difference in the applied distance measure (Jaccard similarity for peptide measurements versus Pearson's correlation for microarray data). The developed metaproteomics approach is an a-priori non-defined approach i.e. identified peptides depend on sample (and not each time the very same peptides in one protein get analysed) whereas the microarray is an a-priori defined approach: i.e. each sample is tested against a fixed set of oligonucleotides.

Individual differences in functions can be very important for the intestinal ecosystem such as the observed large variations in butyrate production or flagellin generation (Fig 5). Notably, one subject (155) had low spectral counts for both butyryl-CoA dehydrogenase and flagellin, both assigned to the Firmicutes. Higher gene counts for flagellar assembly in combination with a higher abundance of members of the phylum Firmicutes were found in a Russian cohort when compared to a Western cohort [43]. The authors concluded that differences in diet i.e. a higher consumption of resistant starch in rural Russian areas, might account for these results. Glucuronate isomerase was among the most abundant 25 functional classes and showed variance between the individuals in this study; an earlier study described that the gene for this

enzyme was present in similar amounts but differed highly in its expression levels across individuals [10].

We detected variations in the levels of an enzyme (sirohydrochlorin cobaltochelatase) involved in the anaerobic synthesis of vitamin B12. An accumulation of the genes involved in B12 cobalamin synthesis has been found in the intestinal microbiota of adults compared to infants [44]. Our data suggest that there might be factors other than age influencing vitamin B12 synthesis.

The observed differences in sugar metabolism, such as the absence of xylose isomerase derived from Actinobacteria in some subjects, may point to a diet lacking the respective substrates. However, as we do not have dietary records of the study participants, causal explanations for the individual as well as temporal differences between the faecal metaproteomes cannot be provided.

We did not detect any obvious functional differences in the faecal metaproteome following the probiotic intervention as compared to the placebo. This is in line with our previous results from the same cohort i.e. the LGG intervention did not change the overall composition of the faecal microbiota; only the amount of *Lactobacilli* (reflecting the ingested strain) differed significantly after the intervention [22]. With LGG representing only up to 0.1% of the total faecal community in the post-intervention samples, it is not surprising that LGG-specific proteins were hardly detected. However, immunological differences between the two groups have been detected [21], evidence that the probiotic product did exert some effects. Rather than colon that represents the predominant source of faecal material, it is more likely that differential protein expression after LGG intake takes place in the ileum which is a region of the gut less densely populated with bacteria and a more active site of signaling. It remains possible that the changes at the protein level in the gut lumen (versus intra-mucosal) are too subtle to be measured with the MS approach used here: i.e. data dependent MS analysis. However, there was a trend that the ratio of bacterial/eukaryotic peptides was increased in the LGG group.

Previously, comparative studies on faecal metaproteomes have been conducted between healthy and diseased subjects. These kind of studies in which an intact and pathologically perturbed intestinal ecosystems have been compared to each other represent much more drastic phenotypic differences than present in our study and may explain the lack of clear differences observed in our cohort [17,18]. In a comparison of the microbial profiles and metaproteomes of healthy and Crohn's Disease (CD) patients, CD patients could be distinguished from the healthy controls based on their 2-DE proteome profiles [18].

## Human-bacterial interactions as assessed by metaproteomics

We were especially interested in intestinal proteins which could serve as a strong indication for the presence of host-microbiota crosstalk. Therefore, we correlated the spectral counts of human and bacteria derived KOs and identified several significant associations. These protein families included the pair of human derived receptor-interacting serine/threonine-protein kinase 4 and the bacterial flagellar hook protein FlgE. This could be one example of a human-microbe interaction in the gut as the human kinase is involved in signalling to elicit an inflammatory response for which flagellins, via TLR5 signalling, could be the stimulus. Human alpha-amylase correlated significantly with the functional families of bacterial putative ABC transport system substrate-binding protein, propionate CoA-transferase, ATP-dependent Clp protease and cytidylate kinase. These correlations could reflect the interdependence regarding nutrient resorption and utilization between host and microbes. Moreover, a negative correlation was observed between the antimicrobial protein DMBT1 (deleted in malignant brain tumors 1 protein) also termed SALSA and known to be present in large amounts in faeces already in early life

[45], and a bacterial, amino-peptidase (bleomycin hydrolase). This finding may point to a physical human-bacterial-protein interaction or to an indirect interaction as agglutination of some other molecule to DMBT1 may suppress the expression of the aminopeptidase.

In addition to the identified correlations, we observed relatively high levels of L-fucose isomerase. As fucose is a typical component of many O- and N-linked glycans produced by mammalian species rather than by plants, its use by the intestinal microbiota is a further indication for a mutualistic interaction between host and microbes. Recently, modulation of the fucosylation of mucus proteins was claimed to act as a physiological mechanism to provide the microbiota with substrates in case of disease [46] and an absence of fucosyl transferase as is the case in non-secretor subjects, was associated with a distinctive intestinal microbiota composition with reduced levels of mucus-degrading bacteria, including *Bifidobacterium* and *Akkermansia* spp. [29]. Here, we could not identify any association between the abundance of L-fucose isomerase and any specific microbial groups.

### Activity of specific microbial groups

The rationale behind studying the metaproteome is that microbial composition and their gene content alone may not be sufficient to explain a certain phenotype. Therefore, we compared the metaproteomic data with the phylogenetic data generated from a microarray. In general, the observations based on the 16S rRNA gene and organism-origin of peptides were often in agreement. However, in some instances, this was not the case and this may reflect the level of biological activity of the specific microbial group. These detected differences between the microbial composition and protein-makeup add further evidence to earlier findings that phylogenetic profiles alone cannot reflect changes in the functions being expressed by the microbiome. In a study where metaproteomic profiles of healthy and CD patients were compared, differences in the amounts of certain proteins were observed although the abundance of corresponding species behaved in a different manner [18].

In this study, we found the amount of Firmicutes in our microarray data to be 1.2 times higher than that inferred from the metaproteomic data, even though the phylogenetic and proteomic measurements for the prominent representative species *F. prausnitzii* were in good agreement. These findings could indicate that Firmicutes proteins may be under-represented in the protein sequence database, so that they get relatively less identified, but *F. prausnitzii* is at least biologically as active as its presence would indicate. In our earlier study, the abundances for Firmicutes were the same but in contrast to the current findings we identified less Bacteroidetes at the peptide level than with the microarray [13]. The difference might be explained by the higher amount of samples. The ratios obtained for Bacteroidetes differ widely between studies partly due to different extraction efficiencies of the DNA extraction protocols being used [32]. The average relative abundance of 10% found here is at the lower end of the range reported for Bacteroidetes [47] but in the same range as in another Finnish cohort [5] and in a Chinese cohort [48]. We cannot exclude the possibility that the “overrepresentation” of Bacteroidetes peptides may reflect their potential underrepresentation in the DNA preparations as in our earlier study where we found relatively less Bacteroidetes peptides compared to the microarray data a different DNA extraction method, i.e. repeated bead beating, was applied. However, as the fluctuations seen at the microarray level were not observed at the peptide level, it might well be that we detected an actual difference in presence and protein production in Bacteroidetes. On the other hand, the observation may also be explained by the fact that extensive DNA transfer has been found in Bacteroidales species leading to a homogenous gene content across species [49] that may facilitate the identification of Bacteroidetes peptides in comparison to the more heterogeneous Firmicutes peptides. As the Bacteroidetes qPCR results

corresponded well to the microarray measurements, a bias due to underrepresentation of certain taxa on the microarray seems unlikely.

The lower relative proportion of Verrucomicrobia in the MS/MS data versus 16S rRNA gene data may be due to an abundance issue, limiting the analysis of proteins of a phylum which contributed only around 1% to the overall microbiota. In an earlier study, we found a high number of Verrucomicrobia peptides but then also the measured DNA reads were well above 1% [15]. The peptide and qPCR results for *Methanobrevibacter* were also in good agreement with each other. The high correlation of the levels of 16S rRNA gene and protein measurement for Actinobacteria gave an indication for an association between the abundance and biological activity of corresponding taxa and also provided an assurance of the reliability of these measurements.

In summary, on one hand, our data allows us to conclude that proteomic data in general reflects phylogenetic composition but on the other hand, it highlights distinct activities and possibly to some extent, analytical bias such as cross-hybridisation and DNA extraction bias. Although today the taxonomic assignment of peptides is possible, due to the far from complete coverage in databases, the assignment is not perfect. Therefore, the technical and annotation issues noted here may explain the results to some extent; this is a limitation of all other metomics studies [50].

## Human proteins in faecal samples—relation to health

Specific human proteins in faeces have been used as markers for certain physiological conditions [51] but from a metaproteomic perspective, differences in gut-derived host proteins have not been studied extensively. Therefore, we also studied the variations in the proteins of the host as the applied approach identifies both microbial and eukaryotic peptides at the same time. We did not fractionate samples to yield a bacterial pellet, as that would selectively detect only those human proteins that are bound to bacterial cells [52]. Since changes of the host and microbiome proteome are being examined at the same time, in theory our approach makes it possible to investigate the mutualistic microbiome host relationship. Overall, the human KOs identified mostly involved in nutrient digestion and the immune response, a result in agreement with earlier metaproteomic studies [13,52]. On average, alpha-amylase, alkaline phosphatase and alpha 1 antitrypsin were the three top human proteins although there were extensive variations between subjects and time points. All of these three proteins have been linked to intestinal or extra-intestinal physiological conditions and they may also be one factor contributing to the individuality of the microbiota as discussed below.

The levels of alpha-amylase may have an effect on the substrate availability for certain microbial species living in the intestine. Recently, alpha-amylase was suggested to stimulate the growth of vaginal *Lactobacillus* spp. [53] by breaking down glycogen. In contrast, *Lactobacillus* strains from baby faeces have been found to exert inhibitory effects upon alpha-amylase and other carbohydrate hydrolases [54]. Therefore, differences in alpha-amylase levels may be reflections of different ecological niches favouring certain microbes. Alpha-amylase may also have a protective effect in the intestine in the same way as has been observed in the mouth where the enzyme inhibits the glucosyltransferase of *Streptococcus mutans* [55]. As observed in this study, the amounts of alpha-amylase have been found to vary with time but also to be consistently non-detectable in some individuals [13]. In a study of patients with diarrhea-type irritable bowel syndrome (IBS-D), elevated levels of tryptic activity and amylase were found [56]. Possibly there are higher levels of amylase during and after diarrhea as the rapid transit of chyme might lead to a decreased (bacterial) degradation of this enzyme. The finding that individual 104 had high levels of alpha-amylase after an episode of diarrhea is support for that

hypothesis. In addition, other factors, such as diet, may have a role, since high alpha-amylase levels were also observed in other individuals who did not report any intestinal symptoms.

Tryptic activity in faecal samples has been associated with intestinal disorders such as IBD [57]. In general, proteinases have been intensively studied in the context of IBD and IBS [58]. The alpha 1 antitrypsin identified here may be involved in regulating the proteinase activities and maintaining a balance of protein degradation in the intestine.

In faeces of mice fed a high fat diet, the levels of several mouse proteins such as alpha 1 antitrypsin were higher in parallel with less identified microbial proteins in comparison to mice consuming a carbohydrate diet [59]. However, the same study also detected higher levels of chymotrypsin B and there was speculation that the proteinase inhibitors may act also through supplementary routes in addition to inhibiting proteinase activity. Therefore, the changes and differences in the human proteins observed here might be due to a certain diet on a specific day [60]. Alpha 1 antitrypsin has been associated with the treatment success of graft-vs.-host disease in GI transplantation in humans [61].

Alkaline phosphatase is a versatile enzyme [62]; in the gut it has been found to combat pathogen invasion into the intestinal epithelium [63]. In a mouse study, the presence of intestinal alkaline phosphatase promoted the growth of a variety of intestinal bacteria [64]. Variations between individuals in the levels of alkaline phosphatase have been detected previously [13]; here very high levels in one individual were found, and high levels of alkaline phosphatase were mostly identified in those subjects who reported intestinal symptoms. Elevated levels of alkaline phosphatase might be one way that compromised intestine attempts to restore equilibrium. It is tempting to speculate that the levels of this enzyme could be a marker of the “quality” (intact, perturbed or somewhere in-between) of the host-microbe interaction.

DMBT1 (deleted in malignant brain tumor) was also one of the top 15 human KOs, its levels were quite abundant in most individuals. As this is a healthy cohort, this result is not unexpected since it has been hypothesized that this protein, which is involved in host defence mechanisms, is indicative of a healthy intestine [52]. Many of the human KOs identified here including alpha 1 antitrypsin, carboxypeptidase A1 and B, and alkaline phosphatase have been found previously to be equally abundant throughout the intestinal tract when proteins at the luminal interface were analysed [65], reflecting their crucial roles in intestinal activity.

Our results suggest that although over time variations in human KOs may be due to pathology, diet and other environmental factors may also contribute to these variations. Inter-individual differences may be influenced by different genetic backgrounds, both human and microbial. It is evident that both human and bacterial proteins need to be studied to fully reveal the mechanisms in the complex ecosystem of the intestine.

## Conclusions

By analysing the proteins from faecal samples of 16 healthy individuals at three measurement time points, the obtained metaproteome was found to contain a functional core but also to show variation across individuals and with time. This study can be viewed as a first step towards defining a healthy baseline, which is essential before one can evaluate the metaproteomes of diseased individuals, similar to the proposal concerning metagenomics [66].

A larger, well-stratified study cohort will be required to address the question of which microbial composition leads to which metaproteomic phenotype, but based on our current data, we conclude that different microbial profiles can lead to differences in the functionality of the intestinal microbiota.

Assigning taxonomies to peptides revealed overall good correlations with the phylogenetic data but caution is required, as presently it is difficult to estimate whether under- and over-



representation of proteomic data is due to the actual biological activity of the microbes or to experimental artefacts such as a (data) analysis bias (technical and annotation bias). Very few studies have compared the composition of the microbiota and its actual functions being carried out, but there is growing evidence that accessing the composition alone is not enough. It appears that while the phylogenetic differences across samples are quite pronounced, metagenomes remain more stable. However, when evaluating the functions being expressed, then the difference may again be larger.

Overall, the inter- and intra-individual differences appeared rather small at the level of COG families. This indicates that with a stable proportion of main functionalities, there is a “healthy balance”, e.g. in protein turnover and replication, in the intestinal ecosystem. Individual differences are more likely to be detected as one moves further down in the functional hierarchy.

This study is one step in the gradual exploration of the biological activity of our gut symbionts. However, before we have a fundamental understanding of the crucial mechanisms present in the intestine, it will be essential to clarify the functions of the many so far unidentified proteins. In addition, functional classification systems have to be optimized for the human microbiome. Overall, metaproteomics is still in its infancy and improvements, especially regarding analytical depth and data analysis, are needed (e.g. in assigning peptides to protein groups for quantitative comparisons) to fill in the large knowledge gaps which are an inherent limitation of functional meta-omic analyses.

The combination of phylogenetic and metaproteomic data helps to gain a better understanding of the intestinal microbiota. For example, it reveals which microbes are living in the intestine and how they contribute to the host’s metabolism including signalling to other bacteria and signalling to the host. In the future, longitudinal studies will be required to clarify the normal flux both in composition and biological activity of our microbes.

## Supporting Information

**S1 Fig. Rarefaction and core analysis of identified bacterial KOs.**  
(TIF)

**S2 Fig. Correlation of human and bacterial KOs (Spearman).**  
(PDF)

**S3 Fig. Hierarchical clustering of KO abundances.**  
(PDF)

**S1 File.** This file contains S4 Fig Hierarchical clustering of oligoprofiles, S5 Fig Distribution of butyrate and propionate synthesis, S6 Fig Clustering of Pearson correlations of bacterial/eukaryotic assignment ratios, S7 Fig A) Distribution of the 5 most abundant phyla for 16S rRNA gene and spectra per sample. B) Bacteroidetes qPCR results, S8 Fig Functionality (COG) across phyla, and S9 Fig Abundance and activity of *F. prausnitzii*.  
(PDF)

**S1 Table. Subject Information.**  
(XLSX)

**S2 Table. Health questionnaire. Supplementary results on metaproteomics findings in combination with health questionnaire information.**  
(PDF)

**S3 Table. Composition of the in-house human intestinal metaproteome database.**  
(DOCX)

**S4 Table. Microarray data on genus-like level.**

(XLSX)

**S5 Table. Overview of mass spectrometric results and taxonomic assignment.**

(XLSX)

**S6 Table. Identified bacterial KOs.**

(XLSX)

**S7 Table. Microbial gut-specific metabolic modules (GMM).**

(CSV)

**S8 Table. Identified human KOs.**

(XLSX)

**S9 Table. Genus-like taxa distribution.**

(TXT)

**S10 Table. Selected KOs of Proteobacteria, Prevotella, Actinobacteria.**

(XLSX)

**S11 Table. Identified KOs from Archaea.**

(XLSX)

## Acknowledgments

Outi Immonen and Nahaison Krips are thanked for technical assistance, the Turku Proteomics Facility of Biocentre Finland at the Turku Centre for Biotechnology for support with LC-MS measurements and data analysis, Jochen Müller for data processing, Ewen Mac Donald for language editing, Katri Korpela for statistical advice, and HITChip team in Wageningen University for phylogenetic measurements.

## Author Contributions

Conceived and designed the experiments: WdV CK AS GC. Performed the experiments: CK. Analyzed the data: CK JS J. Ritari WdV AS J. Raes GF SVS. Contributed reagents/materials/analysis tools: WdV GC RAK AP MdB. Wrote the paper: CK WdV MdB AS JS GC AP J. Ritari J. Raes SVS.

## References

1. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, et al. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol.* 2014; 32: 834–841. doi: [10.1038/nbt.2942](https://doi.org/10.1038/nbt.2942) PMID: [24997786](https://pubmed.ncbi.nlm.nih.gov/24997786/)
2. Karlsson FH, Nookaew I, Nielsen J. Metagenomic data utilization and analysis (MEDUSA) and construction of a global gut microbial gene catalogue. *PLoS Comput Biol.* 2014; 10: e1003706. doi: [10.1371/journal.pcbi.1003706](https://doi.org/10.1371/journal.pcbi.1003706) PMID: [25010449](https://pubmed.ncbi.nlm.nih.gov/25010449/)
3. de Vos WM, de Vos EAJ. Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutr Rev.* 2012; 70: S45–S56. doi: [10.1111/j.1753-4887.2012.00505.x](https://doi.org/10.1111/j.1753-4887.2012.00505.x) PMID: [22861807](https://pubmed.ncbi.nlm.nih.gov/22861807/)
4. Walsh CJ, Guinane CM, O'Toole PW, Cotter PD. Beneficial modulation of the gut microbiota. *FEBS Lett.* 2014; 588: 4120–4130. doi: [10.1016/j.febslet.2014.03.035](https://doi.org/10.1016/j.febslet.2014.03.035) PMID: [24681100](https://pubmed.ncbi.nlm.nih.gov/24681100/)
5. Jalanka-Tuovinen J, Salonen A, Nikkilä J, Immonen O, Kekkonen R, Lahti L, et al. Intestinal microbiota in healthy adults: Temporal analysis reveals individual and common core and relation to intestinal symptoms. *PLoS One.* 2011; 6: e23035. doi: [10.1371/journal.pone.0023035](https://doi.org/10.1371/journal.pone.0023035) PMID: [21829582](https://pubmed.ncbi.nlm.nih.gov/21829582/)
6. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, et al. Temporal variability is a personalized feature of the human microbiome. *Genome Biol.* 2014; 15: 531. doi: [10.1186/s13059-014-0531-y](https://doi.org/10.1186/s13059-014-0531-y) PMID: [25517225](https://pubmed.ncbi.nlm.nih.gov/25517225/)

7. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009; 457: 480–484. doi: [10.1038/nature07540](https://doi.org/10.1038/nature07540) PMID: [19043404](https://pubmed.ncbi.nlm.nih.gov/19043404/)
8. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012; 486: 207–214. doi: [10.1038/nature11234](https://doi.org/10.1038/nature11234) PMID: [22699609](https://pubmed.ncbi.nlm.nih.gov/22699609/)
9. McNulty NP, Yatsunenko T, Hsiao A, Faith JJ, Muegge BD, Goodman AL, et al. The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci Transl Med*. 2011; 3: 106ra106. doi: [10.1126/scitranslmed.3002701](https://doi.org/10.1126/scitranslmed.3002701) PMID: [22030749](https://pubmed.ncbi.nlm.nih.gov/22030749/)
10. Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, et al. Relating the metatranscriptome and metagenome of the human gut. *Proc Natl Acad Sci U S A*. 2014; 111: E2329–38. doi: [10.1073/pnas.1319284111](https://doi.org/10.1073/pnas.1319284111) PMID: [24843156](https://pubmed.ncbi.nlm.nih.gov/24843156/)
11. Zoetendal EG, Raes J, van den Bogert B, Arumugam M, Booiijink CC, Troost FJ, et al. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*. 2012; 6: 1415–1426. doi: [10.1038/ismej.2011.212](https://doi.org/10.1038/ismej.2011.212) PMID: [22258098](https://pubmed.ncbi.nlm.nih.gov/22258098/)
12. Kolmeder CA, de Vos WM. Metaproteomics of our microbiome—developing insight in function and activity in man and model systems. *J Proteomics*. 2014; 97: 3–16. doi: [10.1016/j.jprot.2013.05.018](https://doi.org/10.1016/j.jprot.2013.05.018) PMID: [23707234](https://pubmed.ncbi.nlm.nih.gov/23707234/)
13. Kolmeder CA, de Been M, Nikkilä J, Ritamo I, Mättö J, Valmu L, et al. Comparative metaproteomics and diversity analysis of human intestinal microbiota testifies for its temporal stability and expression of core functions. *PLoS One*. 2012; 7: e29913. doi: [10.1371/journal.pone.0029913](https://doi.org/10.1371/journal.pone.0029913) PMID: [22279554](https://pubmed.ncbi.nlm.nih.gov/22279554/)
14. Cantarel BL, Erickson AR, Verberkmoes NC, Erickson BK, Carey PA, Pan C, et al. Strategies for metagenomic-guided whole-community proteomics of complex microbial environments. *PLoS One*. 2011; 6: e27173. doi: [10.1371/journal.pone.0027173](https://doi.org/10.1371/journal.pone.0027173) PMID: [22132090](https://pubmed.ncbi.nlm.nih.gov/22132090/)
15. Rooijers K, Kolmeder C, Juste C, Doré J, de Been M, Boeren S, et al. An Iterative Workflow for Mining the Human Intestinal Metaproteome. *BMC Genomics*. 2011; 12: 6. doi: [10.1186/1471-2164-12-6](https://doi.org/10.1186/1471-2164-12-6) PMID: [21208423](https://pubmed.ncbi.nlm.nih.gov/21208423/)
16. Kolmeder CA, Ritari J, Verdam FJ, Muth T, Keskitalo S, Varjosalo M, et al. Colonic metaproteomic signatures of active bacteria and the host in obesity. *Proteomics*. 2015; 15: 3544–3552. doi: [10.1002/pmic.201500049](https://doi.org/10.1002/pmic.201500049) PMID: [26255997](https://pubmed.ncbi.nlm.nih.gov/26255997/)
17. Erickson AR, Cantarel BL, Lamendella R, Darzi Y, Mongodin EF, Pan C, et al. Integrated Metagenomics/Metaproteomics Reveals Human Host-Microbiota Signatures of Crohn's Disease. *PLoS One*. 2012; 7: e49138. doi: [10.1371/journal.pone.0049138](https://doi.org/10.1371/journal.pone.0049138) PMID: [23209564](https://pubmed.ncbi.nlm.nih.gov/23209564/)
18. Juste C, Kreil DP, Beauvallet C, Guillot A, Vaca S, Carapito C, et al. Bacterial protein signals are associated with Crohn's disease. *Gut*. 2014; 63: 1566–1577. doi: [10.1136/gutjnl-2012-303786](https://doi.org/10.1136/gutjnl-2012-303786) PMID: [24436141](https://pubmed.ncbi.nlm.nih.gov/24436141/)
19. Pérez-Cobas AE, Gosalbes MJ, Friedrichs A, Knecht H, Artacho A, Eismann K, et al. Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut*. 2013; 62: 1591–1601. doi: [10.1136/gutjnl-2012-303184](https://doi.org/10.1136/gutjnl-2012-303184) PMID: [23236009](https://pubmed.ncbi.nlm.nih.gov/23236009/)
20. Kekkonen RA, Sysi-Aho M, Seppanen-Laakso T, Julkunen I, Vapaatalo H, Oresic M, et al. Effect of probiotic *Lactobacillus rhamnosus* GG intervention on global serum lipidomic profiles in healthy adults. *World J Gastroenterol*. 2008; 14: 3188–3194. PMID: [18506924](https://pubmed.ncbi.nlm.nih.gov/18506924/)
21. Kekkonen RA, Lummela N, Karjalainen H, Latvala S, Tynkkynen S, Järvenpää S, et al. Probiotic intervention has strain-specific anti-inflammatory effects in healthy adults. *World J Gastroenterol*. 2008; 14: 2029–2036. PMID: [18395902](https://pubmed.ncbi.nlm.nih.gov/18395902/)
22. Lahti L, Salonen A, Kekkonen RA, Salojärvi J, Jalanka-Tuovinen J, Palva A, et al. Associations between the human intestinal microbiota, *Lactobacillus rhamnosus* GG and serum lipids indicated by integrated analysis of high-throughput profiling data. *PeerJ*. 2013; 1: e32. doi: [10.7717/peerj.32](https://doi.org/10.7717/peerj.32) PMID: [23638368](https://pubmed.ncbi.nlm.nih.gov/23638368/)
23. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*. 2006; 1: 2856–2860. PMID: [17406544](https://pubmed.ncbi.nlm.nih.gov/17406544/)
24. Huson DH, Weber N. Microbial community analysis using MEGAN. *Methods Enzymol*. 2013; 531: 465–485. doi: [10.1016/B978-0-12-407863-5.00021-6](https://doi.org/10.1016/B978-0-12-407863-5.00021-6) PMID: [24060133](https://pubmed.ncbi.nlm.nih.gov/24060133/)
25. Mesuere B, Devreese B, Debyser G, Aerts M, Vandamme P, Dawyndt P. Unipept: Tryptic Peptide-Based Biodiversity Analysis of Metaproteome Samples. *J Proteome Res*. 2012; 11: 5773–5780. doi: [10.1021/pr300576s](https://doi.org/10.1021/pr300576s) PMID: [23153116](https://pubmed.ncbi.nlm.nih.gov/23153116/)
26. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, et al. Richness of human gut microbiome correlates with metabolic markers. *Nature*. 2013; 500: 541–546. doi: [10.1038/nature12506](https://doi.org/10.1038/nature12506) PMID: [23985870](https://pubmed.ncbi.nlm.nih.gov/23985870/)
27. Ahlroos T, Tynkkynen S. Quantitative strain-specific detection of *Lactobacillus rhamnosus* GG in human faecal samples by real-time PCR. *J Appl Microbiol*. 2009; 106: 506–514. doi: [10.1111/j.1365-2672.2008.04018.x](https://doi.org/10.1111/j.1365-2672.2008.04018.x) PMID: [19200317](https://pubmed.ncbi.nlm.nih.gov/19200317/)

28. Rajilić-Stojanović M, Heilig HG, Molenaar D, Kajander K, Surakka A, Smidt H, et al. Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ Microbiol*. 2009; 11: 1736–1751. doi: [10.1111/j.1462-2920.2009.01900.x](https://doi.org/10.1111/j.1462-2920.2009.01900.x) PMID: [19508560](https://pubmed.ncbi.nlm.nih.gov/19508560/)
29. Wacklin P, Tuimala J, Nikkilä J, Tims S, Mäkituokko H, Alakulppi N, et al. Faecal microbiota composition in adults is associated with the FUT2 Gene Determining the Secretor Status. *PloS One*. 2014; 9: e94863. doi: [10.1371/journal.pone.0094863](https://doi.org/10.1371/journal.pone.0094863) PMID: [24733310](https://pubmed.ncbi.nlm.nih.gov/24733310/)
30. Lahti L, Torrente A, Elo LL, Brazma A, Rung J. A fully scalable online pre-processing algorithm for short oligonucleotide microarray atlases. *Nucleic Acids Res*. 2013; 41: e110. doi: [10.1093/nar/gkt229](https://doi.org/10.1093/nar/gkt229) PMID: [23563154](https://pubmed.ncbi.nlm.nih.gov/23563154/)
31. Rinttilä T, Kassinen A, Malinen E, Krogus L, Palva A. Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. *J Appl Microbiol*. 2004; 97: 1166–1177. PMID: [15546407](https://pubmed.ncbi.nlm.nih.gov/15546407/)
32. Salonen A, Nikkilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M, Kekkonen RA, et al. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J Microbiol Methods*. 2010; 81: 127–134. doi: [10.1016/j.mimet.2010.02.007](https://doi.org/10.1016/j.mimet.2010.02.007) PMID: [20171997](https://pubmed.ncbi.nlm.nih.gov/20171997/)
33. R-core team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: <http://www.R-project.org/>.
34. Levandowsky M, Winter D. Distance between sets. *Nature*. 1971; 234: 34–35.
35. Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics*. 2005; 4: 1419–1440. PMID: [16009968](https://pubmed.ncbi.nlm.nih.gov/16009968/)
36. Lahti L, Salojärvi J, Salonen A, Scheffer M, de Vos WM. Tipping Elements in the Human Intestinal Ecosystem. *Nat Commun*. 2014; 5: 4344. doi: [10.1038/ncomms5344](https://doi.org/10.1038/ncomms5344) PMID: [25003530](https://pubmed.ncbi.nlm.nih.gov/25003530/)
37. Laue H, Smits TH, Schumacher UK, Claros MC, Hartemink R, Cook AM. Identification of *Bifidobacterium wadsworthia* by specific PCR which targets the taurine: pyruvate aminotransferase gene. *FEMS Microbiol Lett*. 2006; 261: 74–79. PMID: [16842362](https://pubmed.ncbi.nlm.nih.gov/16842362/)
38. Huxtable R. Physiological actions of taurine. *Physiol Rev*. 1992; 72: 101–163. PMID: [1731369](https://pubmed.ncbi.nlm.nih.gov/1731369/)
39. Wilmes P, Heintz-Buschart A, Bond PL. A decade of metaproteomics: Where we stand and what the future holds. *Proteomics*. 2015; 15: 3409–3417. doi: [10.1002/pmic.201500183](https://doi.org/10.1002/pmic.201500183) PMID: [26315987](https://pubmed.ncbi.nlm.nih.gov/26315987/)
40. Selvendran RR. The plant cell wall as a source of dietary fiber: chemistry and structure. *Am J Clin Nutr (USA)*. 1984; 39: 320–337.
41. Kamal-Eldin A, Laerke HN, Knudsen KE, Lampi AM, Piironen V, Adlercreutz H, et al. Physical, microscopic and chemical characterisation of industrial rye and wheat brans from the Nordic countries. *Food Nutr Res*. 2009; 53: doi: [10.3402/fnr.v53i0.1912](https://doi.org/10.3402/fnr.v53i0.1912)
42. Lappi J, Salojärvi J, Kolehmainen M, Mykkänen H, Poutanen K, de Vos WM, et al. Intake of whole-grain and fiber-rich rye bread versus refined wheat bread does not differentiate intestinal microbiota composition in Finnish adults with metabolic syndrome. *J Nutr*. 2013; 143: 648–655. doi: [10.3945/jn.112.172668](https://doi.org/10.3945/jn.112.172668) PMID: [23514765](https://pubmed.ncbi.nlm.nih.gov/23514765/)
43. Tyakht AV, Kostyukova ES, Popenko AS, Belenikin MS, Pavlenko AV, Larin AK, et al. Human gut microbiota community structures in urban and rural populations in Russia. *Nat Commun*. 2013; 4: 2469. doi: [10.1038/ncomms3469](https://doi.org/10.1038/ncomms3469) PMID: [24036685](https://pubmed.ncbi.nlm.nih.gov/24036685/)
44. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012; 486: 222–227. doi: [10.1038/nature11053](https://doi.org/10.1038/nature11053) PMID: [22699611](https://pubmed.ncbi.nlm.nih.gov/22699611/)
45. Reichhardt MP, Jarva H, de Been M, Miguel Rodriguez J, Quintana Jimenez E, Loimaranta V, et al. The complement regulating protein SALSA in early life: diverse roles in amniotic fluid and intestine. *J Immunol*. 2014. 193: 5240–5248.
46. Pickard JM, Maurice CF, Kinnebrew MA, Abt MC, Schenten D, Golovkina TV, et al. Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness. *Nature*. 2014; 514: 638–641. doi: [10.1038/nature13823](https://doi.org/10.1038/nature13823) PMID: [25274297](https://pubmed.ncbi.nlm.nih.gov/25274297/)
47. Finucane MM, Sharpston TJ, Laurent TJ, Pollard KS. A taxonomic signature of obesity in the microbiome? Getting to the guts of the matter. *PloS One*. 2014; 9: e84689. doi: [10.1371/journal.pone.0084689](https://doi.org/10.1371/journal.pone.0084689) PMID: [24416266](https://pubmed.ncbi.nlm.nih.gov/24416266/)
48. Ling Z, Liu X, Luo Y, Yuan L, Nelson KE, Wang Y, et al. Pyrosequencing analysis of the human microbiota of healthy Chinese undergraduates. *BMC Genomics*. 2013; 14: 390. doi: [10.1186/1471-2164-14-390](https://doi.org/10.1186/1471-2164-14-390) PMID: [23758874](https://pubmed.ncbi.nlm.nih.gov/23758874/)

49. Coyne MJ, Zitomersky NL, McGuire AM, Earl AM, Comstock LE. Evidence of extensive DNA transfer between bacteroidales species within the human gut. *MBio*. 2014; 5: e01305–14. doi: [10.1128/mBio.01305-14](https://doi.org/10.1128/mBio.01305-14) PMID: [24939888](https://pubmed.ncbi.nlm.nih.gov/24939888/)
50. Lozupone CA, Stombaugh J, Gonzalez A, Ackermann G, Wendel D, Vazquez-Baeza Y, et al. Meta-analyses of studies of the human microbiota. *Genome Res*. 2013; 23: 1704–1714. doi: [10.1101/gr.151803.112](https://doi.org/10.1101/gr.151803.112) PMID: [23861384](https://pubmed.ncbi.nlm.nih.gov/23861384/)
51. Foell D, Wittkowski H, Roth J. Monitoring disease activity by stool analyses: from occult blood to molecular markers of intestinal inflammation and damage. *Gut*. 2009; 58: 859–868. doi: [10.1136/gut.2008.170019](https://doi.org/10.1136/gut.2008.170019) PMID: [19136508](https://pubmed.ncbi.nlm.nih.gov/19136508/)
52. Verberkmoes NC, Russell AL, Shah M, Godzik A, Rosenquist M, Halfvarson J, et al. Shotgun metaproteomics of the human distal gut microbiota. *ISME J*. 2009; 3: 179–189. doi: [10.1038/ismej.2008.108](https://doi.org/10.1038/ismej.2008.108) PMID: [18971961](https://pubmed.ncbi.nlm.nih.gov/18971961/)
53. Spear GT, French AL, Gilbert D, Zariffard MR, Mirmonsef P, Sullivan TH, et al. Human alpha-amylase present in lower-genital-tract mucosal fluid processes glycogen to support vaginal colonization by *Lactobacillus*. *J Infect Dis*. 2014; 210: 1019–1028. doi: [10.1093/infdis/jiu231](https://doi.org/10.1093/infdis/jiu231) PMID: [24737800](https://pubmed.ncbi.nlm.nih.gov/24737800/)
54. Panwar H, Calderwood D, Grant IR, Grover S, Green BD. *Lactobacillus* strains isolated from infant faeces possess potent inhibitory activity against intestinal alpha- and beta-glucosidases suggesting anti-diabetic potential. *Eur J Nutr*. 2014; 53: 1465–1474. doi: [10.1007/s00394-013-0649-9](https://doi.org/10.1007/s00394-013-0649-9) PMID: [24414142](https://pubmed.ncbi.nlm.nih.gov/24414142/)
55. Jespersgaard C, Hajishengallis G, Russell MW, Michalek SM. Identification and characterization of a nonimmunoglobulin factor in human saliva that inhibits *Streptococcus mutans* glucosyltransferase. *Infect Immun*. 2002; 70: 1136–1142. PMID: [11854193](https://pubmed.ncbi.nlm.nih.gov/11854193/)
56. Tooth D, Garsed K, Singh G, Marciani L, Lam C, Fordham I, et al. Characterisation of faecal protease activity in irritable bowel syndrome with diarrhoea: origin and effect of gut transit. *Gut*. 2014; 63: 753–760. doi: [10.1136/gutjnl-2012-304042](https://doi.org/10.1136/gutjnl-2012-304042) PMID: [23911555](https://pubmed.ncbi.nlm.nih.gov/23911555/)
57. Midtvedt T, Zabarovsky E, Norin E, Bark J, Gizatullin R, Kashuba V, et al. Increase of faecal tryptic activity relates to changes in the intestinal microbiome: analysis of Crohn's disease with a multidisciplinary platform. *PloS One*. 2013; 8: e66074. doi: [10.1371/journal.pone.0066074](https://doi.org/10.1371/journal.pone.0066074) PMID: [23840402](https://pubmed.ncbi.nlm.nih.gov/23840402/)
58. Sałaga M, Sobczak M, Fichna J. Inhibition of proteases as a novel therapeutic strategy in the treatment of metabolic, inflammatory and functional diseases of the gastrointestinal tract. *Drug Discov Today*. 2013; 18: 708–715. doi: [10.1016/j.drudis.2013.03.004](https://doi.org/10.1016/j.drudis.2013.03.004) PMID: [23567293](https://pubmed.ncbi.nlm.nih.gov/23567293/)
59. Daniel H, Gholami AM, Berry D, Desmarchelier C, Hahne H, Loh G, et al. High-fat diet alters gut microbiota physiology in mice. *ISME J*. 2014; 8: 295–308. doi: [10.1038/ismej.2013.155](https://doi.org/10.1038/ismej.2013.155) PMID: [24030595](https://pubmed.ncbi.nlm.nih.gov/24030595/)
60. McNulty NP, Wu M, Erickson AR, Pan C, Erickson BK, Martens EC, et al. Effects of diet on resource utilization by a model human gut microbiota containing *Bacteroides cellulosilyticus* WH2, a symbiont with an extensive glycobiome. *PLoS Biol*. 2013; 11: e1001637. doi: [10.1371/journal.pbio.1001637](https://doi.org/10.1371/journal.pbio.1001637) PMID: [23976882](https://pubmed.ncbi.nlm.nih.gov/23976882/)
61. Rodriguez-Otero P, Porcher R, Peffault de Latour R, Contreras M, Bouhnik Y, Xhaard A, et al. Fecal calprotectin and alpha-1 antitrypsin predict severity and response to corticosteroids in gastrointestinal graft-versus-host disease. *Blood*. 2012; 119: 5909–5917. doi: [10.1182/blood-2011-12-397968](https://doi.org/10.1182/blood-2011-12-397968) PMID: [22555971](https://pubmed.ncbi.nlm.nih.gov/22555971/)
62. Lallès J. Intestinal alkaline phosphatase: novel functions and protective effects. *Nutr Rev*. 2014; 72: 82–94. doi: [10.1111/nure.12082](https://doi.org/10.1111/nure.12082) PMID: [24506153](https://pubmed.ncbi.nlm.nih.gov/24506153/)
63. Goldberg RF, Austen WG Jr, Zhang X, Munene G, Mostafa G, Biswas S, et al. Intestinal alkaline phosphatase is a gut mucosal defense factor maintained by enteral nutrition. *Proc Natl Acad Sci U S A*. 2008; 105: 3551–3556. doi: [10.1073/pnas.0712140105](https://doi.org/10.1073/pnas.0712140105) PMID: [18292227](https://pubmed.ncbi.nlm.nih.gov/18292227/)
64. Malo MS, Moaven O, Muhammad N, Biswas B, Alam SN, Economopoulos KP, et al. Intestinal alkaline phosphatase promotes gut bacterial growth by reducing the concentration of luminal nucleotide triphosphates. *Am J Physiol Gastrointest Liver Physiol*. 2014; 306: G826–38. doi: [10.1152/ajpgi.00357.2013](https://doi.org/10.1152/ajpgi.00357.2013) PMID: [24722905](https://pubmed.ncbi.nlm.nih.gov/24722905/)
65. Li X, Leblanc J, Truong A, Vuthoori R, Chen SS, Lustgarten JL, et al. A metaproteomic approach to study human-microbial ecosystems at the mucosal luminal interface. *PLoS One*. 2011; 6: e26542. doi: [10.1371/journal.pone.0026542](https://doi.org/10.1371/journal.pone.0026542) PMID: [22132074](https://pubmed.ncbi.nlm.nih.gov/22132074/)
66. Klymiuk I, Högenauer C, Halwachs B, Thallinger GG, Fricke WF, Steininger C. A Physicians' Wish List for the Clinical Application of Intestinal Metagenomics. *PLoS Med*. 2014; 11: e1001627. doi: [10.1371/journal.pmed.1001627](https://doi.org/10.1371/journal.pmed.1001627) PMID: [24736271](https://pubmed.ncbi.nlm.nih.gov/24736271/)